Express Mail No.: ET318110716US
Date of Deposit: June 25, 2002

Attorney Docket No. 19904- 009
BREESE-9

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANTS:	Calas *et al.*		
ASSIGNEE:	SYNT:EM		
SERIAL NUMBER:	09/484,571	EXAMINER:	Kam, Chih Min
FILING DATE:	June 9, 2000	ART UNIT:	1653
FOR:	LINEAR PEPTIDES DERIVED FROM ANTIBIOTIC PEPTIDES, PREPARATION AND USE FOR VECTORING ACTIVE SUBSTANCES		

June 25, 2002
Boston, Massachusetts

Commissioner for Patents
Washington, D.C. 20231

VERIFIED STATEMENT

For the Examiner's convenience, a substitute and corrected specification has been submitted in the above-referenced application. The substitute specification incorporates all of amendments and corrections to the specification that are set forth in the accompanying Amendment and Response. The amendments/corrections to the specification are entirely cosmetic and/or are submitted to incorporate proper headings, etc. in accordance with 37 CFR 1.77(b). The undersigned, a registered patent attorney, asserts that the substitute specification contains no new matter.

Respectfully submitted,

Dated: June 25, 2002

Barry J. Marenberg, Reg. No. 40,715
Attorneys for Applicants
c/o MINTZ LEVIN
One Financial Center
Boston, Massachusetts 02111
Telephone: (617) 542-6000
Telefax: (617) 542-2241

LINEAR PEPTIDES DERIVED FROM ANTIBIOTIC PEPTIDES, PREPARATION AND USE FOR VECTORING ACTIVE SUBSTANCES

The invention concerns linear peptides derived from antibiotic peptides and their use for vectoring active substances. More particularly, the subject of the invention is new compounds formed from a linear derivative of an antibiotic peptide coupled to at least one active substance, and the preparation of these compounds and compositions containing them.

In addition to their immunity system responsible for specific defence mechanisms against infectious agents, vertebrates have numerous peptides with anti-microbial activity (Nicolas P. et al., 1995, Annual Rev. Microbiol. 49, 277-304). These peptides only exist in invertebrates having a short lifetime and a high renewal rate, in whom a memory immunity system, long in forming and developing appropriate response, is ill-adapted.

The anti-microbial peptides of vertebrates, irrespective of their origin, lower or higher vertebrates, myeloid or non-myeloid tissue, have a certain number of properties in common :

- high basicity due to the presence of numerous arginines and lysines,

- the ability to form amphipathic structures. By amphipathic structure is meant structures in which the hydrophobic residues are separated in space from hydrophilic residues,

- a very wide activity spectrum. They are able to rapidly destroy bacteria (Gram$^+$ and Gram$^-$), fungi, a few protozoa, membrane viruses and even some cancer cell lines.

According to their structure, antibiotic peptides can be divided into three major families :

- amphipathic α-helical antibiotic peptides: cecropins and maganins (Maloy W.L. et al., 1995, BioPolymer 37, 105-122),

- β-stranded antibiotic peptides linked by disulphide bonds : defensins (Lehrer R.I. et al., 1991, Cell 64:229-230 ; Lehrer R.I. et al., 1993, Ann. Rev. Immunol. 11:105-128), protegrins (Kokryakov V.N. et al., 1993, FEBS 337:231-236), tachyplesins (Nakamura T. et al., 1988, J. Biol. Chem. 263:16709-16713 ; Miyata T. et al., 1989, J. Biochem. 106:663-668),

- antibiotic peptides having destructured chains with many angles due to the presence of multiple prolines : bactenecins and PR39 (Frank R.W. et al., 1991, Eur. J. Biochem. 202, 849-854).

Despite the diversity of their sequences, most antibiotic peptides act by direct lysis of the membrane of pathogenic cells. Their basicity promotes their interaction with negatively charged phospholipids, and being amphipathic they are subsequently able to incorporate themselves into the membrane in which they aggregate to form pores through which the cell loses its substance. It is generally accepted that their preferential selectivity for prokaryote cells is due to the special composition of their membranes which contain more anionic phospholipids than those of eukaryotes. Also, the plasma membranes of mammalian cells all contain cholesterol whose role is to modulate their fluidity, which could hinder the incorporation of antibiotic peptides. However, the specificity of the latter for micro-organisms is low, meaning that they show strong cytotoxicity which limits their use.

The presence of antibiotic peptides in vertebrates, and more particularly in mammalians, raises numerous queries. Immunologists assume that the compounds having non-specific anti-microbial activity found in

invertebrates constitute an ancestral means of defence which later developed leading to much more complex memory systems. What is the advantage therefore, in mammalians for example, of having preserved some peptides with antibiotic activity ? It is supposed that these small molecules that are always present in biological fluids, or sequestered in some lymphocyte structures, could form a first line of defence while awaiting the secretion of specific antibodies (Nicolas P. et al., 1995, Annual Rev. Microbiol. 49, 277-304). They could also, within the macrophages, take part in the destruction of plasma membranes of pathogenic organisms.

Regardless of their exact role, antibiotic peptides are of considerable interest owing to their wide spectrum of activity and the difficulty encountered by micro-organisms to set up inactivation strategies. On this account very numerous research studies have been conducted to endeavour to find new molecules and to obtain better performing analogues than the parent peptides. It is possible that in the future these antibiotic peptides are called upon to replace the antibiotics derived from bacteria or fungi. For example, PCT international patent applications published under numbers WO95/03325, WO96/37508 and WO97/02287 describe a new class of antibiotic peptides called "protegrins", isolated from porcine leukocytes or even prepared by chemical synthesis or genetic engineering and having antibacterial, antiviral and antifungal activities.

At the present time, β-stranded antibiotic peptides linked by disulphide bonds (defensins, protegrins, tachyplesins) are a particular subject of research on account of their powerful anti-microbial activity (bacteria, some viruses, fungi and parasites). Within this family, protegrins and tachyplesins are certainly the most promising molecules given the

simplicity of their structure and the relative ease with which they can be synthesised.

The name protegrins denotes a group of five peptides called PG-1, PG-2, PG-3, PG-4 and PG-5 whose sequences are given below, closely resembling and isolated from porcine leukocytes (V.N. Kokryakov et al., FEBS lett. 327, 231-236) :

PG-1 : Arg Gly Gly Arg Leu Cys Tyr Cys Arg Arg Arg Phe Cys Val Cys Val Gly Arg $-NH_2$ (SEQ ID NO : 1)

PG-2 : Arg Gly Gly Arg Leu Cys Tyr Cys Arg Arg Arg Phe Cys Ile Cys Val $-NH_2$ (SEQ ID NO : 2)

PG-3 : Arg Gly Gly Gly Leu Cys Tyr Cys Arg Arg Arg Phe Cys Val Cys Val Gly Arg $-NH_2$ (SEQ ID NO : 3)

PG-4 : Arg Gly Gly Arg Leu Cys Tyr Cys Arg Gly Trp Ile Cys Phe Cys Val Gly Arg $-NH_2$ (SEQ ID NO : 4)

PG-5 : Arg Gly Gly Arg Leu Cys Tyr Cys Arg Pro Arg Phe Cys Val Cys Val Gly Arg $-NH_2$ (SEQ ID NO : 5)

Tachyplesins (Tamura H. et al., 1993, Chem. Pharm. Bul. Tokyo 41, 978-980) denoted T1, T2 and T3 and polyphemusins (Muta T., 1994, CIBA Found. Sym. 186, 160-174) denoted P1 and P2 whose sequences are given below, are homologous peptides isolated from the hemolymph of two crabs *Tachyplesus tridentatus* for Tachyplesins T1, T2 and T3, and *Limmulus polyphemus* for Polyphemusins P1 and P2.

P1 : Arg Arg Trp Cys Phe Arg Val Cys Tyr Arg Gly Phe Cys Tyr Arg Lys Cys Arg $-NH_2$ (SEQ ID NO : 6)

P2 : Arg Arg Trp Cys Phe Arg Val Cys Tyr Lys Gly Phe Cys Tyr Arg Lys Cys Arg $-NH_2$ (SEQ ID NO : 7)

T1 : Lys Trp Cys Phe Arg Val Cys Tyr Arg Gly Ile Cys Tyr Arg Arg Cys Arg $-NH_2$ (SEQ ID NO : 8)

T2 : Arg Trp Cys Phe Arg Val Cys Tyr Arg Gly Ile Cys Tyr Arg Lys Cys Arg $-NH_2$ (SEQ ID NO : 9)

T3 : Lys Trp Cys Phe Arg Val Cys Tyr Arg Gly Ile Cys Tyr Lys Arg Cys Arg -NH_2 (SEQ ID NO : 10)

Protegrins, tachyplesins and polyphemusins contain a high proportion of base residues (lysines and arginines) and have four cysteines which form two parallel disulphide bonds. These three families of peptides also show homologies with some defensins in particular with the human defensin NP-1 (Kokryakov V.N. et al., 1993, Febs Let. 327, 231-236).

Tachyplesins and protegrins have a closely resembling three-dimensional structure. It is an anti-parallel β strand stabilised by the two disulphide bonds. These bonds play an important role in the antibacterial activity of protegrins and tachyplesins. Their removal, either by protecting the SH groups with acetamidomethyls, or by replacing the cysteines with alanines or glycines, leads to obtaining analogues virtually devoid of *in vivo* activity (Lehrer R.I. et al., 1996, Eur. J. Biochem. 240:352-357).

As previously indicated, protegrins and tachyplesins have substantial lysis activity on prokaryote cells. Research work conducted by the Applicant on the cytotoxicity of these peptides on cultured mammalian cells, have shown that, prior to the death of the cells, there are non-negligible quantities of protegrins and tachyplesins in the cytoplasm of said cells. It was considered that the presence of peptides in the cytoplasm could be the outcome of transport via pores, but these pores are only permeable to ions and small molecules and their diameter is too small to give passageway to antibiotic peptides. It would seem that protegrins and tachyplesins, in addition to perforating the plasma membrane, are able to pass through it.

The cytotoxicity and antimicrobial activity of protegrins and tachyplesins are known to derive from

their ability to aggregate inside the membrane to form multimeric channels (Mangoni M. et al., 1996, Febs Let. 383, 93-98). The Applicant therefore considered that this aggregation might be connected with the tertiary structure of these antibiotic peptides, which comprise several cysteine residues, and linear derivatives of protegrins and tachyplesins in which the cysteines are replaced by various natural amino acids have been prepared. These peptides were coupled, at their N-terminal end, to a fluorescent molecule or to biotin and the distribution of these markers inside the cell was observed under confocal microscopy.

In this way, it was found that these peptides are non-toxic and have no lytic activity but are, on the other hand, able to pass rapidly through the membranes of mammalian cells via a passive mechanism.

These linear derivatives of antibiotic peptides therefore constitute a new, non-toxic, system for vectoring active substances.

By vectoring system is meant, according to the invention, a process capable of conveying said active substance to a target, such as for example :

- to cause an active substance to pass through the cell membrane and to allow the distribution of said substance in the cytoplasm and/or in the nuclear compartment.

- to bring an active substance to a particular organ, for example to cause this active substance to pass through the blood-brain barrier,

- to force this active substance to interact specifically with a given cell type, erythrocytes for example.

The subject of the present invention is therefore peptides derived from antibiotic peptides or analogues

thereof, characterised in that they are devoid of a disulphide bond.

By analogue of antibiotic peptides is meant a peptide whose amino acid sequence has been modified without causing any modification in the antibiotic properties of said peptide.

The absence of a disulphide bond in the peptides of the invention, may be obtained by any means known to those skilled in the art, for example by :

- removing, or replacing with other amino acids, the cysteine residues of the antibiotic peptide sequence,

- blocking the -SH groups of the cysteine residues such that they do not form a disulphide bond,

provided, evidently, that the peptide obtained has vectoring properties that are not toxic for the previously described cells.

These modifications may be conducted during the preparation of the peptides of the invention, more particularly by chemical synthesis or the expression of a gene coding for said peptide, or directly on an antibiotic peptide through the action of chemical agents enabling the opening and blocking of the -SH groups of the cysteine residues.

The above modifications advantageously concern all the cysteine residues of the antibiotic peptide, but should the presence of a single cysteine residue not allow the formation of a disulphide bond, the peptides of the invention may contain a single cysteine. Natural antibiotic peptides generally have 4 or 6 cysteine residues able to form two or three disulphide bonds, therefore in the peptides of the invention only one of these cysteines can be maintained and the three or five others are modified or blocked.

The antibiotic peptides from which the peptides of the invention are derived may be defensins, protegrins, tachyplesins or their analogues whose antibiotic properties are imparted to them by their tertiary structure resulting from the presence of disulphide bonds.

Linear peptides of the invention meet one of the following formulas :

Baa Xaa Xaa Baa Xaa Xaa Xaa Xaa Baa Baa Baa Xaa Xaa Xaa Xaa Xaa Xaa Baa (I) (SEQ ID NO : 11)

Baa Baa Xaa Xaa Xaa Baa Xaa Xaa Xaa Baa Xaa Xaa Xaa Xaa Baa Baa Xaa Baa (II) (SEQ ID NO : 12)

which may also be represented by the following single formula (III) :

Baa (Xaa Baa) Xaa (Xaa Baa) Xaa (Xaa Baa) Xaa Xaa (Xaa Baa) Baa (Xaa Baa)Xaa Xaa Xaa (Xaa Baa) (Xaa Baa) Xaa Baa (SEQ ID NO : 38)

in which :

- the Baa groups, identical or different, represent an amino acid residue whose side chain carries a base group, and

- the Xaa groups, identical or different, represent an aliphatic or aromatic amino acid residue,

or are made up of a sequence of at least 5, preferably at least 7, successive amino acids of either of formulas (I) or (II), if this sequence has vectoring properties that are non-toxic for the previously described cells.

Baa and Xaa may or may not be natural amino acids, including D-amino acids.

As an example the following denotations of Baa and Xaa may be cited :

- Baa is chosen from among arginine, lysine, diaminoacetic acid, diaminobutyric acid, diaminoproprionic acid, ornithine.

- Xaa is chosen from among glycine, alanine, valine, norleucine, isoleucine, leucine, cysteine, cysteine^{Acm}, penicillamine, methionine, serine, threonine, asparagine, glutamine, phenylalanine, histidine, tryptophan, tyrosine, proline, Amino butyric acid, carboxylic amino-1-cyclohexane acid, Amino isobutyric acid, carboxylic 2-aminotetraline, 4-bromophenylalanine, tert-Leucine, 4-chlorophenylalanine, β-cyclohexylalanine, 3,4-dichlorophenylalanine, 4-fluorophenylalanine, homoleucine, β-homoleucine, homophenylalanine, 4-methylphenylalanine, 1-naphthylalanine, 2-naphthylalanine, 4-nitrophenylalanine, 3-nitrotyrosine, norvaline, phenylglycine, 3-pyridylalanine, β-(2-Thienyl)-alanine.

The invention also concerns peptide derivatives having the formula (I) or (II) such as said peptides in retro form, or moieties of peptides having the formula (I) or (II) made up of five, preferably seven, successive amino acids of either formula (I) or (II).

Among the peptides of the invention, special mention can be made of those meeting the following formulas :

Arg Xaa Xaa Arg Xaa Uaa Xaa Uaa Arg Arg Arg Xaa Uaa Xaa Uaa Xaa Xaa Arg -NH_2 (V) (SEQ ID NO : 13)

Arg Arg Xaa Uaa Xaa Arg Xaa Uaa Xaa Arg Xaa Xaa Uaa Xaa Arg Arg Uaa Arg -NH_2 (VI) (SEQ ID NO : 14)

in which :

- Uaa represents serine or threonine

- the Xaa groups, identical or different, represent an amino acid which may or may not be natural

(including D-amino acids), either aliphatic or aromatic, such as among glycine, alanine, valine, norleucine, isoleucine, leucine, cysteine, cysteine^{Acm}, penicillamine, methionine, serine, threonine, asparagine, glutamine, phenylalanine, histidine, tryptophan, tyrosine, proline, Amino butyric acid, carboxylic amino-1-cyclohexane acid, Amino isobutyric acid, carboxylic 2-aminotetraline, 4-bromophenylalanine, tert-Leucine, 4-chlorophenylalanine, β-cyclohexylalanine, 3,4-dichlorophenylalanine, 4-fluorophenylalanine, homoleucine, β-homoleucine, homophenylalanine, 4-methylphenylalanine, 1-naphthylalanine, 2-naphthylalanine, 4-nitrophenylalanine, 3-nitrotyrosine, norvaline, phenylglycine, 3-pyridylalanine, β-(2-Thienyl)-alanine.

Among the peptides of formulas (I) and (II) or their derivatives, the invention specifically considers those derived from protegrins and tachyplesins referred to in tables I and II below.

Table I : Protegrin derivatives

Code	Sequence	Modification
SM1738	Arg Gly Gly Arg Leu Ser Tyr Ser Arg Arg Arg Phe Ser Val Ser Val Gly Arg (SEQ ID NO : 15)	Head of series
SM1736	Arg Gly Gly Arg Leu Ser Tyr Ser Arg Arg Arg Phe Ser Val Ser Val Gly Arg (SEQ ID NO : 16)	Aa of D form of SM1738
SM1727	Arg Gly Val Ser Val Ser Phe Arg Arg Arg Ser Tyr Ser Leu Arg Gly Gly Arg (SEQ ID NO : 17)	Retro form of SM1738
SM1739	Glu Gly Gly Glu Leu Ser Tyr Ser Glu Glu Glu Phe Ser Val Ser Val Gly Glu (SEQ ID NO : 18)	Reversed charge (R → E)
SM2187	Arg Gly Gly Arg Leu Ala Tyr Arg Leu Leu Arg Phe Ala Ile Arg Val Gly Arg (SEQ ID NO : 19)	Increased amphipathicity
SM2188	Oaa Gly Gly Oaa Xaa Xaa Baa Oaa Xaa Xaa Oaa Baa Xaa Xaa Xaa Oaa Xaa Gly (SEQ ID NO : 20)	Increased hydrophobicity
SM2189	Arg Ala Ala Arg Leu Gly Tyr Arg Xaa Xaa Arg Phe Gly Zaa Arg Val Gly Arg (SEQ ID NO : 21)	Increased amphipathicity
SM2194	Tyr Arg Arg Arg Phe Ser Val Ser Val Arg (SEQ ID NO : 22)	C-terminal end of SM2193
SM2195	Arg Arg Leu Ser Tyr Ser Arg Arg Arg Phe (SEQ ID NO : 23)	N-terminal end of SM2193

SM2193	Arg Arg Leu Ser Tyr Ser Arg Arg Arg Phe Ser Val Ser Val Arg (SEQ ID NO : 24)	Reduced flexibility (G deletion)
SM2196	Arg Gly Gly Arg Leu Ser Tyr Ser Arg Arg Arg Phe Ser Thr Ser Thr Gly Arg (SEQ ID NO : 25)	Inhibition dimerisation

Table II : Tachyplesin derivatives

Code	Sequence	Modification
SM1726	Lys Trp Ser Phe Arg Val Ser Tyr Arg Gly Ile Ser Tyr Arg Arg Ser Arg (SEQ ID NO : 26)	Head of series
SM2307	Arg Trp Ser Phe Arg Val Ser Tyr Arg Gly Ile Ser Tyr Arg Arg Ser Arg (SEQ ID NO : 27)	K → R mutation
SM2392	Arg Trp Ser Phe Arg Val Ser Tyr Arg Gly Ile Ser Tyr Arg Arg Ser Arg (SEQ ID NO : 28)	Aa of D form (of SM2307)
SM2309	Lys Trp Ser Phe Arg Val Ser Tyr Arg Gly Ile Ser Tyr Arg Arg Ser Arg (SEQ ID NO : 29)	Aa of D form (of SM1726)
SM2310	Arg Ser Arg Arg Tyr Ser Ile Gly Arg Tyr Ser Val Arg Phe Ser Trp Lys (SEQ ID NO : 30)	Retro form

SM2190	Oaa Baa Xaa Baa Oaa Xaa Xaa Baa Oaa Gly Xaa Oaa Baa Xaa Xaa Oaa Xaa (SEQ ID NO : 31)	Increased hydrophobicity
SM2191	Lys Trp Ala Phe Arg Val Ala Tyr Arg Gly Ile Arg Tyr Leu Leu Arg Leu (SEQ ID NO : 32)	Increased amphipathicity
SM2192	Lys Tyr Ala Trp Arg Val Ala His Arg Gly Ile Arg Trp Leu Leu Arg Xaa (SEQ ID NO : 33)	Increased amphipathicity

In the sequences of tables I and II above, Baa represents Naphthylalanine, Oaa represents Ornithine, Xaa represents Norleucine and Zaa represents Norvaline.

The invention also concerns the use of the above peptides for vectoring one or more active substances both for therapeutic and for diagnostic applications. As active substance, the invention gives particular consideration to proteins or moieties of proteins, such as polypeptides or peptides, antibodies or parts of antibodies, nucleic acids and oligonucleotides or ribozymes, or even, obviously, active chemical molecules for the treatment or prevention of human or animal pathologies, such as for example, but not restricted to, anti-tumorals, antivirals, anti-inflammatories, agents preventing the degradation of organs and/or tissues, etc...

In the field of diagnostics, the active substance may be a radioactive marker, a stain marker, or any other means or substance able to reveal a metabolism or a pathology.

A further purpose of the invention is therefore compounds of formula (IV) as follows and compositions containing them :

$(Y)_n$——(A)——$(Z)_m$—— IV

in which :

- A represents a linear peptide derived from an antibiotic peptide in accordance with the invention,
- Z represents an active substance, such as defined above,
- Y represents a signal agent,
- n is 0 or more, advantageously 0 or 1,
- m is 1 or more, preferably up to 10, advantageously up to 5.

Therefore, the above formula (IV) compounds are formed from a peptide of the invention coupled with one or more active substances, identical or different, represented by the (Z) group in formula (IV), and optionally one or more signal agents, represented by the (Y) group in formula (IV), having an addressing role for the compound of formula (IV) towards a cell type, a site or cell compartment or a given tissue. More particularly, the signal agent (Y) is an oligopeptide or a protein, such as a signal peptide, a nuclear localising signal, an antibody moiety, or a chemical molecule ligand or anti-ligand of a receptor.

In a special embodiment of the compounds of formula (IV), group (Y) is fixed to group (Z).

This coupling, symbolised by the horizontal lines in formula (IV), may be conducted by any acceptable linking means, taking into consideration the chemical nature, the size and number of groups (Z) and (Y) in the compounds of formula (IV), such as covalent, hydrophobic or ionic bonds, which may or may not be cleaved in physiological media. Coupling may be conducted at any site of peptide (A), at which functional groups such as -

OH, -SH, -COOH, -NH2 are naturally present or have been inserted.

The invention gives consideration to the fixation of several (Z) groups to one and the same site of peptide (A) either directly, if this site comprises several functional groups as is the case for a C- or N-terminal lysine, or indirectly via an intermediate group carrying several reaction groups enabling the fixation of several (Z) groups.

The preferred coupling positions for the active substance are at the N-terminal and C-terminal ends or at the primary amino groups carried by the side chains of the lysines of peptide (A). If the C-terminal end of peptide (A) is used to attach active substance (Z), the N-terminal end is available for optional coupling to a signal agent (Y) enabling the compound of the invention to be addressed either towards the nucleus, or towards a given tissue type.

For example, if the C-terminal end of a linear peptide of the invention is used to couple an active substance made up of a fluorescent marker, or biotin, or a medicinal molecule such as doxorubicin, the covalent peptide-drug complex distributes itself after administration within the cytoplasm of the target cell. It is possible to bring this complex into the nuclear compartment by using the N-terminal end of the peptide to couple a short basic sequence, for example of around 7 amino acids, corresponding to a nuclear localising signal. Under these conditions, the biotin or doxorubicin are found in the cell nucleus.

In the same way, it is possible to vector a drug towards a given cell type, by using the N-terminal end of the linear peptide of the invention coupled at its C-terminal end to a medicinal agent, to add a peptide sequence able to specifically recognise a determinant

present on the surface of cell type. Synthetic pentadecapeptide αM2 for example (Swolapenko G.B. et al., 1995, The Lancet 346, 1662-65) a moiety of a monoclonal antibody, directed against an antigen expressed by breast cancer cells (Tumour Associated Antigen Polymorphic Epithelial Mucin) maintains good affinity for these cells. It is therefore possible, by associating αM2 with a linear peptide-medicinal agent complex, to bring this group preferably towards the cells which express the antigen characteristic related to breast cancer.

The compounds of formula (IV) may be prepared by chemical synthesis or by using molecular biology techniques.

For chemical syntheses, commercially available equipment can be used allowing the incorporation of non-natural amino acids, such as D enantiomers and residues with side chains of different hydrophobicity and size to those of their natural homologues. At the time of synthesis it is evidently possible to conduct a wide range of modifications, for example to insert a lipid (prenyl or myristyl) on the N-terminal so as to be able to anchor the peptide of the invention and hence the formula (IV) compound to a lipid membrane such as that of a liposome made up of positively charged lipids. It is also possible to replace one or more peptide bonds (-CO-NH-) by equivalent structures such as $-CO-N(CH_3)-$, $-CH_2-CH_2-$, $-CO-CH_2-$, or to interpose groups such as $-CH_2-$, -NH-, -O-.

It is also possible to obtain formula (IV) compounds, or part thereof having a protein nature, from an encoding nucleic acid sequence. A further purpose of the invention is a nucleic acid molecule comprising or made up of a nucleic sequence coding for a linear peptide derived from an antibiotic peptide. More particularly, the invention concerns a nucleic acid molecule comprising

at least one sequence coding for a formula (IV) compound or part thereof having a protein nature. These nucleic acid sequences may be DNAs or RNAs and be associated with control sequences and/or inserted in vectors. The vector used is chosen in relation to the host to which it will be transferred ; it may be any vector such as a plasmid. These nucleic acids and vectors are useful for producing the linear peptides and formula (IV) compounds, or part of the latter having a protein nature, in a host cell. The preparation of these vectors and the production or expression in a host of linear peptides or formula (IV) compounds may be conducted using molecular biology and genetic engineering techniques well known to those skilled in the art.

By way of example, said method for producing a peptide of the invention consists of :

- transferring a nucleic acid molecule or a vector containing said molecule into a host cell,

- culturing said host cell under conditions enabling the production of the peptide,

- isolating, by any appropriate means, the peptides of the invention.

The host cell used in this type of method may be chosen from among prokaryotes or eukaryotes, in particular from among bacteria, yeasts, mammalian, plant or insect cells. The invention therefore also concerns transformed cells expressing the linear peptides or formula (IV) compounds or part of the latter having a protein nature.

The invention also relates to :

- pharmaceutical compositions comprising as active ingredient at least one formula (IV) compound optionally associated with an acceptable vehicle or carrier,

- diagnostic agents containing at least one formula (IV) compound.

Other characteristics and advantages of the invention will become apparent in the following description concerning the preparation of formula (IV) compounds and the research work which led to revealing the vectoring properties of the linear peptides of the invention derived from antibiotic peptides.

Example 1 : Fixing biotin and doxorubicin onto linear analogues of antibiotic peptides.

1) Preparation of linear peptides

The three peptides with the sequences given below were synthesised :

Arg Gly Gly Arg Leu Xaa Tyr Xaa Arg Arg Arg Phe Xaa Val Xaa Val Gly Arg-NH_2 (SEQ ID NO : 34)

Arg Arg Trp Xaa Phe Arg Val Xaa Tyr Arg Gly Phe Xaa Tyr Arg Lys Xaa Arg-NH_2 (SEQ ID NO : 35)

Lys Trp Xaa Phe Arg Val Xaa Tyr Arg Gly Ile Xaa Tyr Arg Arg Xaa Arg-NH_2 (SEQ ID NO : 36)

in which Xaa represents the serine, threonine or alanine residues.

These peptides are respectively derived from the sequences of Protegrin PG-1 having the formula :

Arg Gly Gly Arg Leu Cys Tyr Cys Arg Arg Arg Phe Cys Val Cys Val Gly Arg -NH_2 (SEQ ID NO : 1)

of Tachyplesin 1 having the formula :

Lys Trp Cys Phe Arg Val Cys Tyr Arg Gly Ile Cys Tyr Arg Arg Cys Arg -NH_2 (SEQ ID NO : 8)

of Polyphemusin having the formula :

Lys Trp Xaa Phe Arg Val Xaa Tyr Arg Gly Ile Xaa Tyr Arg Arg Xaa Arg-NH_2 (SEQ ID NO : 36)

These three peptides may be prepared indifferently either from BOC chemistry or from FMOC chemistry using conventional synthesis methods in solid or homogeneous phase.

2) Fixing biotin onto linear peptides

The peptide is synthesised in solid phase and, after incorporation of the N-terminal arginine, 5-aminopentanoic acid is added. The FMOC or BOC N-terminal is removed, and on the peptide still adhering to the resin, the N-hydroxy succimido biotin ester is caused to react in dimethylformamide. After 15 hours' reaction at room temperature, the biotinylated peptide is cut from the carrier through the action of trifluoroacetic acid or hydrofluoric acid following well-established protocols in peptide chemistry. The peptide is then purified by high pressure liquid chromatography.

3) Fixing doxorubicin onto a linear peptide

To fix doxorubicin, solid phase synthesis is made of the peptide having the formula :

Arg Gly Gly Arg Leu Xaa Tyr Xaa Arg Arg Arg Phe Xaa Val Xaa Val Gly Arg-NH_2 (SEQ ID NO : 37)

After cleaving from the purification substrate, the peptide is treated with glutaric anhydride in the presence of triethylamine. The peptide is then purified and the -COOH group carried by the glutaryl at the N-terminal is activated by the diisopropylcarbodiimide and 1-hydroxybenzotriazole mixture. After two hours' reaction at room temperature, the doxorubicin is added and the mixture is stirred for 12 hours at 0°C. The peptide-doxorubicin unit is then purified by high pressure liquid chromatography.

Example 2 : Ability of the linear peptides of the invention to pass through cell membranes.

1) Cell models

The ability of the peptides to pass through the membranes was tested on various cell types (MCF7, MCF7R, HL60, HL60R, HeLa).

The cells are cultured on RPMI 1640 (Gibco) to which is added 10 % (v/v) fetal calf serum, 2mM glutamine and 2mM pencillin/streptomycin at 37°C. 30 000 cells are seeded in Lab Tek chambers and cultured for 1 day.

2) Treatment with linear peptides-biotin prepared according to example 1 (2)

The cells are incubated in Opti-Mem (Gibco) for one hour before being treated for variable time periods with biotin-labelled peptides.

The latter are obtained in accordance with example 1 (2) by treating 1 equivalent of linear peptide with 2 equivalents of N-hydroxysuccinimide biotin ester, then purified by high pressure liquid chromatography.

The cells are then fixed with a 3.7 % solution of paraformaldehyde for 5 minutes at 25°C, then rinsed three times with PBS. They are then permeabilised with 0.1 % Triton (1 min. room temperature). After three rinsings in PBS the cells are incubated 10 min with 200 µl TexRed antibodies diluted to 300^{th} and rinsed three times in PBS. The slides are finally mounted with a Mowiol-Dabco solution and observed under an Axiophot photomicroscope.

3) Treatment with linear peptides-doxorubicin prepared in accordance with example 1 (3)

The cells are incubated for 15 minutes, then rinsed with PBS and the doxorubicin present in the cell is determined by chromatography.

4) Results

a) Among the peptides studied, those which pass the most easily through the membranes are those with the following formulas :

Arg Xaa Xaa Arg Xaa Uaa Xaa Uaa Arg Arg Arg Xaa Uaa Xaa Uaa Xaa Xaa Arg -NH_2 (V) (SEQ ID NO : 13)

Arg Arg Xaa Uaa Xaa Arg Xaa Uaa Xaa Arg Xaa Xaa Uaa Xaa Arg Arg Uaa Arg -NH_2 (VI) (SEQ ID NO : 14)

in which

- Uaa represents serine or threonine, and

- the Xaa groups, identical or different, represent an amino acid which may or may not be natural (including D-amino acids), either aliphatic or aromatic, such as glycine, alanine, valine, norleucine, isoleucine, leucine, cysteine, cysteine^{Acm}, penicillamine, methionine, serine, threonine, asparagine, glutamine, phenylalanine, histidine, tryptophan, tyrosine, proline, Amino butyric acid, carboxylic amino-1-cyclohexane acid, Amino isobutyric acid, carboxylic 2-aminotetraline, 4-bromophenylalanine, tert-Leucine, 4-chlorophenylalanine, β-cyclohexylalanine, 3,4-dichlorophenylalanine, 4-fluorophenylalanine, homoleucine, β-homoleucine, homophenylalanine, 4-methylphenylalanine, 1-naphthylalanine, 2-naphthylalanine, 4-nitrophenylalanine, 3-nitrotyrosine, norvaline, phenylglycine, 3-pyridylalanine, β-(2-Thienyl)-alanine..

b) The results of the experiments conducted with doxorubicin show a significant increase in the plasma and nuclear concentration of doxorubicin when the latter is coupled with the linear peptide of the invention compared with the use of doxorubicin alone.

c) The experiments with biotin were conducted more especially on MCF7 cells treated at different times with a complex of biotin and a peptide of the invention having the formula :

biotin-Arg Gly Gly Arg Leu Ser Tyr Ser Arg Arg Arg Phe Ser Val Ser Val Gly Arg-NH_2 (SEQ ID NO : 15)

This work was photographed (not shown) :

- Control in which the cell was treated with biotin alone,

- Treatment of the cell for 2 minutes with a complex : biotin-linear peptide of the invention,

- Treatment of the cell for 30 minutes with a complex: biotin-linear peptide of the invention.

It can be seen in these photographs that biotin alone does not enter the cell and accumulates weakly around the cell. Conversely, with the complex of the invention, it can be seen that the biotin is rapidly led by the linear peptide of the invention inside the cell in which it is present in the cytoplasm and cell nucleus.

Example 3 : Internalisation ability of the linear peptides of the invention

Linear peptides of the invention derived from Protegrins and Tachyplesins were tested on different cell lines for the purposes of assessing their respective internalisation.

1) Experimental conditions

The cells were seeded at approximately 10^4 cells per dish, 24 h before the addition of biotinylated peptides. On the day of the experiment confluence was 60-80%. The biotinylated peptides are incubated with the cells at a concentration of 10 µM for 15 minutes at 37°C in an atmosphere of 95% humidity and 5% CO2 in an OptiMem medium. The cells are washed three times with PBS at room temperature and are then fixed with formalin (3.7% formaldehyde in PBS, 10 min at room temperature). They are then washed in PBS and permeabilised for 15 min with PBS-TritonX-100. Development is made with streptavidin-Texas-Red for 15 min away from light and the cells are then slide mounted. They are observed under a fluorescence microscope and compared with a positive

control (Ap43-58), well described in the literature, and with a negative control.

The cell nuclei were Hoechst stained.

2) Cell lines

All the lines tested are of human origin and were commercially obtained from ATCC.

- Non-tumoral lines : MRC5 (lung fibroblast), HuVeC (endothelial, umbilical cord)

- Tumoral lines : HT29 (colon carcinoma), HepG2 (hepatoblastoma), A172 (glioblastoma), HMCB (melanoma).

The cells are cultured at 37°C in an atmosphere of 95% humidity and 5% CO2. The culture medium is the one recommended by ATCC.

3) Tested peptides

The two series of tested peptides are those given in tables I and II.

4) Results

The internalisation results are shown in tables III and IV below. The peptides penetrate the cells with different degrees of internalisation. Some (such as SM1739 and SM2190) are not internalised whereas others (such as SM2307, SM2187 etc...) penetrate with good efficacy. We also observed that some peptides enter into a given cell type more than in others. SM2196 for example has better internalisation in tumoral cells (HepG2, A172 and HT29) than in non-tumoral cells (MRC5 and HuVeC). Conversely, the SM1738 peptide has greater penetration in non-tumoral lines than in tumoral lines. These results suggest the existence of cell tropism.

Generally it would appear that the retro form of the heads of series does not significantly modify internalisation. Increased hydrophobicity has a negative effect for both families of tested peptides. It is

therefore advisable to avoid increasing hydrophobia. On the other hand, an increase in amphipathicity seems to have a positive effect at least for the Protegrin family.

Table III : Protegrin derivatives

	HepG2	A172	HMCB	HuVeC	MRC5	HT29	Internalisation
SM1738	+	+	+	+++	+++	+	Reference
SM1727	0	++	++	+++	+	+	No significant effect
SM1736	++	+	+++	++++	++++	+	No significant effect
SM1739	0	+	+	0	0	0	Negative effect
SM2187	+++	+++	++++	+++	++++	+++	Positive effect
SM2189	+++	++	+++	++	++++	++	Positive effect
SM2188	0	0	0	++	0	0	Negative effect
SM2193	++	++	+++	++	0	0	Negative effect
SM2194	0	+	+++	+	+	0	Negative effect
SM2195	++++	0	+++	+	+	++++	Contradictory
SM2196	++++	++++	++	+	+	++++	Tropism

Fluorescence microscopy photographs of internalisation are shown in figures 1 and 2. In the A172 and HT29 lines, the SM1738 peptide, shown as an example, appears to be mainly localised in the cytoplasm and in a perinuclear zone. For the HuVec line, the peptide is mainly localised in the cytoplasm. The left column corresponds to nucleus staining with Hoechst.

Table IV : Tachyplesin derivatives

	HepG2	A172	HMCB	HuVeC	MRC5	HT29	Internalisation
SM1726	+++	+	++++ +	+++	+++	+++	Reference
SM2310	ND	++	++++	+++	++	+++	No effect
SM2309	ND	++++	++	++	++++	++++	ND
SM2191	++	++	++	ND	+++	+++	No effect

SM2192	+	+++	++++	+++	++++	++	No effect
SM2190	0	0	0	0	0	0	Negative effect
SM2307	ND	++++ +	++++ +	++++	++++	++++ +	Positive effect
SM2392	ND	+++	++++	++	+++	++++	No effect

ND = not determined

The internalisation photographs are shown in appended figures 3 and 4. For the 3 cell lines shown (A172, HT29, HuVeC) the biotinylated peptide is localised in the cytoplasm in diffuse manner and also distinctly labels the nucleolus. The left column corresponds to nucleus staining with Hoechst.

Example 4 : Internalisation of vectored doxorubicin

The cells are seeded to approximately 10^4 cells per dish 24 h before the addition of the products. On the day of the experiment confluence is 60-80%. The free doxorubicin or the doxorubicin coupled to the SM1738 vector are incubated with the MCF7 cells at a concentration of 10 µM for 60 minutes at 37°C in an atmosphere of 95% humidity and 5% CO2 in the culture medium. The subcell localisation of doxorubicin, naturally fluorescent, was determined by confocal microscopy. The results are given in appended figure 5. The localisation is partly cytoplasmic and partly nuclear. The nucleus in this case is labelled in diffuse manner.

In the peptide sequences listed below, the amino acids are represented by their one-letter code, but they may also be represented by their three-letter code according to the following nomenclature :

A Ala alanine

C	Cys	cysteine
D	Asp	aspartic acid
E	Glu	glutamic acid
F	Phe	phenylalanine
G	Gly	glycine
H	His	histidine
I	Ile	isoleucine
K	Lys	lysine
L	Leu	leucine
M	Met	methionine
N	Asn	asparagine
P	Pro	proliné
Q	Gln	glutamine
R	Arg	arginine
S	Ser	serine
T	Thr	threonine
V	Val	valine
W	Trp	tryptophan
Y	Tyr	tyrosine